PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/705, C12Q 1/68, C07K 16/28, A61K 31/70, 35/00, 38/00, 38/17, 48/00, G01N 33/53, 33/68

A1

(11) International Publication Number: **WO 98/36062**

(43) International Publication Date: 20 August 1998 (20.08.98)

(21) International Application Number: PCT/GB98/00434

(22) International Filing Date: 12 February 1998 (12.02.98)

(30) Priority Data:
9703011.8 13 February 1997 (13.02.97) GB
97305485.1 22 July 1997 (22.07.97) EP
(34) Countries for which the regional or international application was filed: AT et al.

(71) Applicant *(for all designated States except US)*: SMITHKLINE BEECHAM PLC [GB/GB]; New Horizons Court, Brentford, Middlesex TW8 9EP (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: TERRETT, Jonathan, Alexander [GB/GB]; SmithKline Beecham Pharmaceuticals, New Frontiers Science Park South, Third Avenue, Harlow, Essex CM19 5AW (GB). KENWRICK, Susan, Jane [GB/GB]; The University of Cambridge, School of Clinical Medicine, Dept. of Medicine, Addenbrookes Hospital, Level 5, P.O. Box 157, Hill Road, Cambridge, Cambridgeshire CB2 2QQ (GB). WANG, Bo [CN/CN]; The University of Cambridge, School of Clinical Medicine, Dept. of Medicine, Addenbrookes Hospital, Level 5, P.O. Box 157, Hill Road, Cambridge, Cambridgeshire CB2 2QQ (GB).

(74) Agent: CONNELL, Anthony, Christopher; SmithKline Beecham, Corporate Intellectual Property, Two New Horizons Court, Brentford, Middlesex TW8 9EP (GB).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: NEURAL CELL ADHESION MOLECULE SPLICING VARIANTS

(57) Abstract

NrCAMvar polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing NrCAMvar polypeptides and polynucleotides in the design of protocols for the treatment of diabetes, obesity and cancer, among others, and diagnostic assays for such conditions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

NEURAL CELL ADHESION MOLECULE SPLICING VARIANTS

FIELD OF INVENTION

This invention relates to newly identified splice-variant polynucleotides, polypeptides encoded by them and to the use of such polynucleotides and polypeptides, and to their production. More particularly, the polynucleotides and polypeptides of the present invention relate to the cell adhesion molecule family, hereinafter referred to as NrCAMvar. The invention also relates to inhibiting or activating the action of such polynucleotides and polypeptides.

BACKGROUND OF THE INVENTION

The NgCAM-related cell adhesion molecule NrCAM, also called bravo, was first identified and characterised in chick by Grumet *et al.*,(1991). Although the sequence of rat NrCAM is not published, it has been cloned and sequenced (Davis and Bennett, 1994). This cell surface glycoprotein is a member of the immunoglobulin (Ig) superfamily, and is very similar in structure to chick NgCAM, human and mouse L1 and chick neurofascin. Each consists of six Ig domains, five fibronectin type III-like repeats, a transmembrane domain and an intracellular region. These neural cell surface proteins play a critical role in nervous system development. Studies from Bennett *et al.*, (Bennett and Gilligan, 1993; Davis and Bennett,1994) suggested that these molecules, including chick and rat NrCAM, have ankyrin binding activity suggesting that they may be important in membrane-cytoskeletal connections in brain. A role for NrCAM in the *in vivo* guidance of chick commissural neurones has been identified and distinguished from that of NgCAM (Stoeckli and Landmesser, 1995). Chick NrCAM in floor plate cells together with axonin-1 on commissural growth cones is essential for accurate pathfinding at the midline whereas NgCAM is required for fasciculation of the commissural neurites. As well as interacting with axonin-1, NrCAM can also bind at the cell surface with F11, another member of the Ig superfamily (Morales *et al.*, 1993).

Recently a highly conserved human homologue to chick NrCAM was described (Lane *et al.*, Genomics 35 (3), 456-465 (1996)) with 82% amino acid identity to the chick protein. The transmembrane and intracellular domains of human NrCAM are 100% identical to the chick homologue while percent identities for individual extracellular domains vary from 66% for IgVI to 93% for IgIV. Lane *et al* . identified two alternatively spliced exons, AE12 encoding a 12-amino-acid section 5' to FNIII-5, and AE93 encoding the 93-amino-acids corresponding to the whole of FNIII-5 (Figure 2). Four different isoforms were found: with both AE12 and AE93, with only AE12 or AE93, and without

either AE12 or AE93. In addition to AE12 and AE93, two more splice variants have been identified in chick, AE19 and AE10. AE19 encodes a 19-amino-acid section between IgII and IgIII while AE10 is a 10 amino-acid section between IgVI and FNIII-1 (Grumet *et al.*, 1991). Using human NrCAM probes, Lane *et al.* observed one major RNA band of ~7.0kb in multiple brain tissues including amygdala, caudate nucleus, corpus callosum, hippocampus, hypothalamus, substantia nigra, subthalamic nucleus, and thalamus. In chick, the same size of RNA was found in brain tissue but not in embryo heart, gizzard or liver on Northern blots.

This indicates that these cell adhesion molecules have an established importance in vertebrate development and are consequently candidates for therapeutic targets. Clearly there is a need for identification and characterization of further members and variants, including splice variants, of the cell adhesion molecule family which can play a role in preventing, ameliorating or correcting dysfunctions or diseases.

SUMMARY OF THE INVENTION

In one aspect, the invention relates to NrCAMvar polypeptides and recombinant materials and methods for their production. Another aspect of the invention relates to methods for using such NrCAMvar polypeptides and polynucleotides. Such uses include the treatment of diabetes, obesity and cancer, among others. In still another aspect, the invention relates to methods to identify agonists and antagonists using the materials provided by the invention, and treating conditions associated with NrCAMvar imbalance with the identified compounds, including diabetes, obesity and cancer. Yet another aspect of the invention relates to diagnostic assays for detecting diseases associated with inappropriate NrCAMvar activity or levels, including diabetes, obesity and cancer.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the nucleotide and deduced amino acid sequence of a human NrCAMvar; SEQ ID NOS: 1 and 2, respectively.

Figure 2 shows a comparison of the sequences of human NrCAMvar of the present invention and human and chick NrCAM cDNAs.

DESCRIPTION OF THE INVENTION

Definitions

The following definitions are provided to facilitate understanding of certain terms used frequently herein.

"NrCAMvar" refers, among others, generally to a polypeptide having the amino acid sequence set forth in SEQ ID NO:2 or an allelic variant thereof.

"NrCAMvar activity or NrCAMvar polypeptide activity" or "biological activity of the NrCAMvar or NrCAMvar polypeptide" refers to the metabolic or physiological function of said NrCAMvar including similar activities or improved activities or these activities with decreased undesirable side-effects. Also included are antigenic and immunogenic activities of said NrCAMvar.

"NrCAMvar gene" refers to a polynucleotide having the nucleotide sequence set forth in SEQ ID NO:1 or allelic variants thereof and/or their complements.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

"Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications have been made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres.

"Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993 and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", *Meth Enzymol* (1990) 182:626-646 and Rattan *et al.*, "Protein Synthesis: Posttranslational Modifications and Aging", *Ann NY Acad Sci* (1992) 663:48-62.

"Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions,

deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Identity" is a measure of the identity of nucleotide sequences or amino acid sequences. In general, the sequences are aligned so that the highest order match is obtained. "Identity" *per se* has an art-recognized meaning and can be calculated using published techniques. See, e.g.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, 1988; BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, 1993; COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, 1987; and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans (Carillo, H., and Lipton, D., *SIAM J Applied Math* (1988) 48:1073). Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H., and Lipton, D., *SIAM J Applied Math* (1988) 48:1073. Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCS program package (Devereux, J., *et al.*, *Nucleic Acids Research* (1984) 12(1):387), BLASTP, BLASTN, FASTA (Atschul, S.F. *et al.*, *J Molec Biol* (1990) 215:403).

The invention discloses a new splice variant of NrCAM (termed NrCAMvar) which comprises the AE10K2 sequence, which is absent in the published human NrCAM sequence (Lane *et al.*, 1996) but which is present in the chick sequence. In addition the NrCAMvar does

not have the AE10K1 sequence which is present in the Lane *et al.* human sequence (Figure 2). NrCAMvar is expressed at high levels in the brain, pancreas and adrenal cortex and at lower levels in placenta, adrenal medulla, thyroid and testis. The published human NrCAM, however, appears not to be expressed in the pancreas.

Polypeptides of the Invention

In one aspect, the present invention relates to novel NrCAMvar polypeptides. The NrCAMvar polypeptides include the polypeptide of SEQ ID NO:2; as well as polypeptides comprising the amino acid sequence of SEQ ID NO: 2. Preferably NrCAMvar polypeptide exhibit at least one biological activity of NrCAMvar.

The NrCAMvar polypeptides may be in the form of the "mature" protein or may be a part of a larger protein such as a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production.

Biologically active fragments of the NrCAMvar polypeptides are also included in the invention. A fragment is a polypeptide having an amino acid sequence that entirely is the same as part, but not all, of the amino acid sequence of the aforementioned NrCAMvar polypeptides. As with NrCAMvar polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region.

Preferred fragments include, for example, truncation polypeptides having the amino acid sequence of NrCAMvar polypeptides, except for deletion of a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Also preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Biologically active fragments are those that mediate NrCAMvar activity, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those that are antigenic or immunogenic in an animal, especially in a human.

Preferably, all of these polypeptide fragments retain the biological activity of the NrCAMvar, including antigenic activity. Variants of the defined sequence and fragments also

form part of the present invention. Preferred variants are those that vary from the referents by conservative amino acid substitutions -- i.e., those that substitute a residue with another of like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr.

The NrCAMvar polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

Polynucleotides of the Invention

Another aspect of the invention relates to NrCAMvar polynucleotides. NrCAMvar polynucleotides include isolated polynucleotides which encode the NrCAMvar polypeptides and fragments, and polynucleotides closely related thereto. More specifically, NrCAMvar polynucleotide of the invention include a polynucleotide comprising the nucleotide sequence set forth in SEQ ID NO:1 encoding a NrCAMvar polypeptide of SEQ ID NO: 2, and polynucleotides having the particular sequence of SEQ ID NO:1. Also included under NrCAMvar polynucleotides are a nucleotide sequence which has sufficient identity to a nucleotide sequence contained in SEQ ID NO:1 to hybridize under conditions useable for amplification or for use as a probe or marker. The invention also provides polynucleotides which are complementary to such NrCAMvar polynucleotides.

NrCAMvar of the invention is structurally related to other proteins of the cell adhesion molecules, as shown by the results of sequencing the cDNA encoding human NrCAMvar. The cDNA sequence contains an open reading frame encoding a polypeptide of 1304 amino acids. Amino acid of sequence of Figure 1 (SEQ ID NO:2) has about >99% identity (using BlastP) in 1299 amino acid residues with Human NrCAM (Lane, RP et al, Genomics 35 (3), 456-465 (1996)). Nucleotide sequence of Figure 1 (SEQ ID NO:1) has about >99% identity (using BlastN) in 3897 nucleotide residues with Human NrCAM (Genomics 35 (3), 456-465 (1996)). Figure 2 shows the splice variant AE10K.

One polynucleotide of the present invention encoding NrCAMvar may be obtained using standard cloning and screening, from a cDNA library derived from mRNA in cells of human adrenal using the expressed sequence tag (EST) analysis (Adams, M.D., *et al. Science* (1991) 252:1651-1656; Adams, M.D. *et al., Nature*, (1992) *355*:632-634; Adams, M.D., *et al., Nature* (1995) 377 Supp:3-174). Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

The nucleotide sequence encoding NrCAMvar polypeptide of SEQ ID NO:2 may be identical over its entire length to the coding sequence set forth in Figure 1 (SEQ ID NO:1), or may be a degenerate form of this nucleotide sequence encoding the polypeptide of SEQ ID NO:2, or may be highly identical to a nucleotide sequence that encodes the polypeptide of SEQ ID NO:2.

When the polynucleotides of the invention are used for the recombinant production of NrCAMvar polypeptide, the polynucleotide may include the coding sequence for the mature polypeptide or a fragment thereof, by itself; the coding sequence for the mature polypeptide or fragment in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc Natl Acad Sci USA* (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Further preferred embodiments are polynucleotides encoding NrCAMvar variants comprise the amino acid sequence NrCAMvar polypeptide of Figure 1 (SEQ ID NO:2) in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acid residues are substituted, deleted or added, in any combination.

The present invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences.

Polynucleotides of the invention, which are identical or sufficiently identical to a nucleotide sequence contained in SEQ ID NO:1, may be used as hybridization probes for cDNA and genomic DNA, to isolate full-length cDNAs and genomic clones encoding NrCAMvar polypeptide and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to the NrCAMvar gene. Such hybridization techniques are known to those of skill in the art. Typically these nucleotide sequences are 70% identical, preferably 80% identical, more preferably 90% identical to that of the referent. The probes generally will comprise at least 15 nucleotides. Preferably, such probes will have at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will range between 30 and 50 nucleotides.

In one embodiment, to obtain a polynucleotide encoding NrCAMvar comprises the steps of screening an appropriate library under stingent hybridization conditions with a labeled probe having the SEQ ID NO: 1 or a fragment thereof, and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to those of skill in the art. Stringent hybridization conditions are as defined above or alternatively conditions under overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

The polynucleotides and polypeptides of the present invention may be employed as research reagents and materials for discovery of treatments and diagnostics to animal and human disease.

Vectors, Host Cells, Expression

The present invention also relates to vectors which comprise a polynucleotide or polynucleotides of the present invention, and host cells which are genetically engineered with vectors of the invention and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis et al., *BASIC METHODS IN MOLECULAR BIOLOGY* (1986) and Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989) such as calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as streptococci, staphylococci, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used. Such systems include, among others, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements,

from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides to produce a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL* (*supra*).

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the desired polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

If the NrCAMvar polypeptide is to be expressed for use in screening assays, generally, it is preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If NrCAMvar polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide; if produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

NrCAMvar polypeptides can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

Diagnostic Assays

This invention also relates to the use of NrCAMvar polynucleotides for use as diagnostic reagents. Detection of a mutated form of NrCAMvar gene associated with a dysfunction will provide a diagnostic tool that can add to or define a diagnosis of a disease or susceptibility to a disease which results from under-expression, over-expression or altered expression of NrCAMvar. Individuals carrying mutations in the NrCAMvar gene may be detected at the DNA level by a variety of techniques.

Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used

directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled NrCAMvar nucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. See, e.g., Myers *et al.*, *Science* (1985) 230:1242. Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method. See Cotton *et al.*, *Proc Natl Acad Sci USA* (1985) 85: 4397-4401. In another embodiment, an array of oligonucleotides probes comprising NrCAMvar nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability. (See for example: M.Chee et al., Science, Vol 274, pp 610-613 (1996)).

The diagnostic assays offer a process for diagnosing or determining a susceptibility to Diabetes, obesity and cancer through detection of mutation in the NrCAMvar gene by the methods described.

In addition, Diabetes, obesity and cancer, can be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of NrCAMvar polypeptide or NrCAMvar mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as an NrCAMvar polypeptide, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Chromosome Assays

The nucleotide sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise

chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

Antibodies

The polypeptides of the invention or their fragments or analogs thereof, or cells expressing them can also be used as immunogens to produce antibodies immunospecific for the NrCAMvar polypeptides. The term "immunospecific" means that the antibodies have substantiall greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

Antibodies generated against the NrCAMvar polypeptides can be obtained by administering the polypeptides or epitope-bearing fragments, analogs or cells to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., *Nature* (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, *Immunology Today* (1983) 4:72) and the EBV-hybridoma technique (Cole *et al.*, MONOCLONAL ANTIBODIES AND CANCER THERAPY, pp. 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can also be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography.

Antibodies against NrCAMvar polypeptides may also be employed to treat Diabetes, obesity and cancer, among others.

Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with NrCAMvar polypeptide, or a fragment thereof, adequate to produce antibody and/or T cell immune response to protect said animal from Diabetes, obesity and cancer, among others. Yet another aspect of the invention relates to a method of inducing immunological response in a mammal which comprises, delivering NrCAMvar polypeptide via a vector directing expression of NrCAMvar polynucleotide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

Further aspect of the invention relates to an immunological/vaccine formulation (composition) which, when introduced into a mammalian host, induces an immunological response in that mammal to a NrCAMvar polypeptide wherein the composition comprises a NrCAMvar polypeptide or NrCAMvar gene. The vaccine formulation may further comprise a suitable carrier. Since NrCAMvar polypeptide may be broken down in the stomach, it is preferably administered parenterally (including subcutaneous, intramuscular, intravenous, intradermal etc. injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation instonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Screening Assays

The NrCAMvar polypeptide of the present invention may be employed in a screening process for compounds which activate (agonists) or inhibit activation of (antagonists, or otherwise called inhibitors) the NrCAMvar polypeptide of the present invention. Thus, polypeptides of the invention may also be used to assess identify agonist or antagonists from, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These agonists or antagonists may be natural substrates, ligands, receptors, etc., as the case may be, of the polypeptide of the present invention; or may be structural or functional mimetics of the polypeptide of the present invention. See Coligan *et al.*, *Current Protocols in Immunology* 1(2):Chapter 5 (1991).

NrCAMvar polypeptides are ubiquitous in the mammalian host and are responsible for many biological functions, including many pathologies. Accordingly, it is desirous to find compounds and drugs which stimulate NrCAMvar polypeptide on the one hand and which can inhibit the function of NrCAMvar polypeptide on the other hand. In general, agonists are employed for therapeutic and prophylactic purposes for such conditions as Diabetes, obesity and cancer. Antagonists may be employed for a variety of therapeutic and prophylactic purposes for such conditions as Diabetes, obesity and cancer.

In general, such screening procedures may involve using appropriate cells which express the NrCAMvar polypeptide or respond to NrCAMvar polypeptide of the present invention. Such cells include cells from mammals, yeast, *Drosophila* or *E. coli*. Cells which express the NrCAMvar polypeptide (or cell membrane containing the expressed polypeptide) or respond to NrCAMvar polypeptide are then contacted with a test compound to observe binding, or stimulation or inhibition of a functional response. The ability of the cells which were contacted with the candidate compounds is compared with the same cells which were not contacted for NrCAMvar activity.

The assays may simply test binding of a candidate compound wherein adherence to the cells bearing the NrCAMvar polypeptide is detected by means of a label directly or indirectly associated with the candidate compound or in an assay involving competition with a labeled competitor. Further, these assays may test whether the candidate compound results in a signal generated by activation of the NrCAMvar polypeptide, using detection systems appropriate to the cells bearing the NrCAMvar polypeptide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Standard methods for conducting such screening assays are well understood in the art.

Examples of potential NrCAMvar polypeptide antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligands, substrates, receptors, etc., as the case may be, of the NrCAMvar polypeptide, e.g., a fragment of the ligands, substrates, receptors, or small molecules which bind to the polypetide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

Prophylactic and Therapeutic Methods

This invention provides methods of treating an abnormal conditions related to both an excess of and insufficient amounts of NrCAMvar polypeptide activity, including diabetes, obesity and cancer.

If the activity of NrCAMvar polypeptide is in excess, several approaches are available. One approach comprises administering to a subject an inhibitor compound (antagonist) as

hereinabove described along with a pharmaceutically acceptable carrier in an amount effective to inhibit activation by blocking binding of ligands to the NrCAMvar polypeptide, or by inhibiting a second signal, and thereby alleviating the abnormal condition.

In another approach, soluble forms of NrCAMvar polypeptides still capable of binding the ligand in competition with endogenous NrCAMvar polypeptide may be administered. Typical embodiments of such competitors comprise fragments of the NrCAMvar polypeptide.

In still another approach, expression of the gene encoding endogenous NrCAMvar polypeptide can be inhibited using expression blocking techniques. Known such techniques involve the use of antisense sequences, either internally generated or separately administered. See, for example, O'Connor, *J Neurochem* (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Alternatively, oligonucleotides which form triple helices with the gene can be supplied. See, for example, Lee *et al., Nucleic Acids Res* (1979) 6:3073; Cooney *et al., Science* (1988) 241:456; Dervan *et al., Science* (1991) 251:1360. These oligomers can be administered *per se* or the relevant oligomers can be expressed *in vivo*.

For treating abnormal conditions related to an under-expression of NrCAMvar and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically effective amount of a compound which activates NrCAMvar polypeptide, i.e., an agonist as described above, in combination with a pharmaceutically acceptable carrier, to thereby alleviate the abnormal condition. Alternatively, gene therapy may be employed to effect the endogenous production of NrCAMvar by the relevant cells in the subject. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo*. For overview of gene therapy, see Chapter 20, *Gene Therapy and other Molecular Genetic-based Therapeutic Approaches*, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996).

Formulation and Administration

Peptides, such as the soluble form of NrCAMvar polypeptides, and agonists and antagonist peptides or small molecules, may be formulated in combination with a suitable pharmaceutical carrier. Such formulations comprise a therapeutically effective amount of the

polypeptide or compound, and a pharmaceutically acceptable carrier or excipient. Such carriers include but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. Formulation should suit the mode of administration, and is well within the skill of the art. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

Preferred forms of systemic administration of the pharmaceutical compositions include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if properly formulated in enteric or encapsulated formulations, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels and the like.

The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100 µg/kg of subject. Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy" as described above. Thus, for example, cells from a subject may be engineered with a polynucleotide, such as a DNA or RNA, to encode a polypeptide *ex vivo*, and for example, by the use of a retroviral plasmid vector. The cells are then introduced into the subject.

Examples

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples illustrate, but do not limit the invention.

Example 1

Cloning of human NrCAMvar cDNA

The HGS EST database was screened using the chick NrCAM sequence and three HGS EST clones (EST99669, EST237133, and EST373834) were obtained. EST99669 and EST237133 clones were from human adrenal cDNA library while EST373834 was from human striatum cDNA library. cDNA clones EST237133 and EST373834 contained several *Eco*RI or *Eco* RI/*Xho* I fragments, suggesting that inserts from several different genes may be present. Only those fragments containing the EST sequence homologous to NrCAM were subcloned and used for further characterisation. These clones were end sequenced and used as probes labelled with [α-^{32}P]dCTP (Amersham) to screen a human fetal brain (20-wk) λgt11 cDNA library (Clontech). Four positive cDNA clones were isolated and the inserts were cloned into pBluescript plasmids (Maniatis *et al.*, 1982). Additional sections of the gene were isolated using gene-specific primers to amplify cDNA from the human fetal brain Marathon^{TM} cDNA (Clontech) and the λgt11 fetal brain cDNA library. All sequencing was performed on an ABI373 sequencer using the ABI PRISM^{TM} dye terminator cycle sequencing ready reaction kit. Sequences were assembled using the Wisconsin GCG package.

Comparison of both DNA and amino acid sequences for human NrCAMvar (SEQ ID NO:1) and the published NrCAM sequence (Lane *et al.*) revealed that they were >99% identical in overlapping regions. The DNA sequence of human NrCAMvar of SEQ ID NO:1 is 77.1% identical to that of chick gene while the amino acid sequence (SEQ ID NO:2) is 80% identical. Evidence for alternative splicing of AE19, AE12 and AE93, was obtained either through sequencing of cDNA clones or PCR products from human cDNA and in addition two novel regions, AE10K and AE10L were found to be differentially absent in the present (AE10K) versus Lanes (AE10L) sequence (see Figure 2).

a) Exon structure:

EST99669 was found to contain a contiguous genomic sequence not homologous to any sequence of the chick gene. Examination of this sequence revealed a splice donor consensus sequence. This probably corresponds to an intron:exon boundary represented in cDNA due to incomplete mRNA processing. A donor and an acceptor were present in EST373834. In addition to alternatively spliced regions AE12 and AE93, identified by Lane *et al.* (1996), AE19 was also found to be absent from some cDNA fragments obtained from human fetal brain cDNA. In addition, two novel alternatively spliced regions, encoding 10-amino-acid sections, were identified (see Figure 2). AE93 was

absent in EST237133 which was from human adrenal cDNA library while both AE12 and AE93 were observed in PCR products from human adult brain cDNA.

b) Results from Northern blots:

A mRNA band of ~7.0kb was observed for human brain, placenta, pancreas, adrenal medulla and cortex, thyroid, and testis tissues. The results also showed that this gene is highly expressed in brain, pancreas, and adrenal cortex tissues (the levels of mRNA on the blots used are controlled at Clontech and samples are tested for their integrity by hybridisation with an actin gene probe).

c) Chromosomal localisation:

To obtain precise localisation of the human NrCAM locus , primers from an intron (sbp12) and from an exon (sbp13) were designed according to the sequence of EST373834. These primers produced a PCR product of 214bp, and were used to screen for the presence of this gene in the Genebridge 4 radiation hybrid panel. The results were analysed by the WIGCR experimental mapping server and are shown in Table 1. These data place the human NrCAM gene on the long arm of chromosome 7 at 7q21-22 between D7S666 and D7S658.

Table 1: Data vectors obtained from testing Genebridge 4 Radiation Hydrids panel using primers sbp12 and 13.

D7S666	00000 00100 10000 00010 00000 10110 00001 10000 00100 00021 10101
	00000 01100 10010 11000 01000 11120 01000 210
NrCAM	00000 00100 10001 00010 00000 10110 00000 00000 01100 00000 10101
	00000 01100 10010 11000 11000 11110 01000 010
D7S658	00000 00100 10000 00011 00000 10110 00001 10000 01100 00001 10101
	00000 01100 10010 11000 01002 11120 01000 210

Each digit corresponds to one of 93 cell lines in the radiation hybrid panel. 0 and 1 represent negative and positive PCR assays respectively. 2 shows that the assay was contradictory between duplicate experiments or was untested.

Since pancreatic function is intimately involved in the development of diabetes, NrCAMvar becomes a target molecule in the management of this disease. This suggestion is supported by the genomic mapping data. A locus for non-insulin-dependent diabetes mellitus (NIDDM), also called Type II diabetes, has been mapped to the same region of chromosome 7 as NrCAMvar (Prochazka, 1995).

SEQUENCE LISTING

INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3997 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCATCGTAAT TCGCCTAATG CAGCTTAAAA TAATGCCGAA AAAGAAGCGC TTATCTGCGG 60
GCAGAGTGCC CCTGATTCTC TTCCTGTGCC AGATGATTAG TGCACTGGAA GTACCTCTTG 120
ATCCAAAACT TCTTGAAGAC TTGGTACAGC CTCCAACCAT CACCCAACAG TCTCCAAAAG 180
ATTACATTAT TGACCCTCGG GAGAATATTG TAATCCAGTG TGAAGCCAAA GGGAAACCGC 240
CCCCAAGCTT TTCCTGGACC CGTAATGGGA CTCATTTTGA CATCGATAAA GACCCTCTGG 300
TCACCATGAA GCCTGGCACA GGAACGCTCA TAATTAACAT CATGAGCGAA GGGAAAGCTG 360
AGACCTATGA AGGAGTCTAT CAGTGTACAG CAAGGAACGA ACGCGGAGCT GCAGTTTCTA 420
ATAACATTGT TGTCCGCCCA TCCAGATCAC CATTGTGGAC CAAAGAAAAA CTTGAACCAA 480
TCACACTTCA AAGTGGTCAG TCTTTAGTAC TTCCCTGCAG ACCCCCAATT GGATTACCAC 540
CACCTATAAT ATTTTGGATG GATAATTCCT TTCAAAGACT TCCACAAAGT GAGAGAGTTT 600
CTCAAGGTTT GAATGGGGAC CTTTATTTTT CCAATGTCCT CCCAGAGGAC ACCCGCGAAG 660
ACTATATCTG TTATGCTAGA TTTAATCATA CTCAAACCAT ACAGCAGAAG CAACCTATTT 720
CTGTGAAGGT GATTTCAGTG GATGAATTGA ATGACACTAT AGCTGCTAAT TTGAGTGACA 780
CTGAGTTTTA TGGTGCTAAA TCAAGTAGAG AGAGGCCACC AACATTTTTA ACTCCAGAAG 840
GCAATGCAAG TAACAAAGAG GAATTAAGAG GAAATGTGCT TTCACTGGAG TGCATTGCAG 900
AAGGACTGCC TACCCCAATT ATTTACTGGG CAAAGGAAGA TGGAATGCTA CCCAAAAACA 960
GGACAGTTTA TAAGAACTTT GAGAAAACCT TGCAGATCAT TCATGTTTCA GAAGCAGACT 1020
CTGGAAATTA CCAATGTATA GCAAAAAATG CATTAGGAGC CATCCACCAT ACCATTTCTG 1080
TTAGAGTTAA AGCGGCTCCA TACTGGATCA CAGCCCCTCA AAATCTTGTG CTGTCCCCAG 1140
GAGAGGATGG GACCTTGATC TGCAGAGCTA ATGGCAACCC CAAACCCAGA ATTAGCTGGT 1200
TAACAAATGG AGTCCCAATA GAAATTGCCC CTGATGACCC CAGCAGAAAA ATAGATGGCG 1260

ATACCATTAT TTTTTCAAAT GTTCAAGAAA GATCAAGTGC AGTATATCAG TGCAATGCCT 1320
CTAATGAATA TGGATATTTA CTGGCAAACG CATTTGTAAA TGTGCTGGCT GAGCCACCAC 1380
GAATCCTCAC ACCTGCAAAC ACACTCTACC AGGTCATTGC AAACAGGCCT GCTTTACTAG 1440
ACTGTGCCTT CTTTGGGTCA CCTCTCCCAA CCATCCAGTG GTTTAAAGGA GCTAAAGGAA 1500
GTGCTCTTCA TGAAGATATT TATGTTTTAC ATGAAAATGG AACTTTGGAA ATTCCTGTGG 1560
CCCAAAAGGA CAGTACAGGA ACTTATACGT GTGTTGCAAG GAATAAATTA GGGATGGCGA 1620
AGAATGAAGT TCACTTAGAA ATCAAAGATC CTACATGGAT CGTTAAACAG CCCGAATATG 1680
CAGTTGTGCA AAGAGGGAGC ATGGTGTCCT TTGAATGCAA AGTGAAACAT GATCACACCT 1740
TATCCCTCAC TGTCCTGTGG CTGAAGGACA ACAGGGAACT GCCCAGTGAT GAAAGGTTCA 1800
CTGTTGACAA GGATCATCTA GTGGTAGCTG ATGTCAGTGA CGATGACAGC GGGACCTACA 1860
CGTGTGTGGC CAACACCACT CTGGACAGCG TCTCCGCCAG CGCTGTGCTT AGCGTTGTTG 1920
CTCCTACTCC AACTCCAGCT CCCGTTTACG ATGTCCCAAA TCCTCCGCTT GACTTAGAAC 1980
TGACAGATCA ACTTGACAAA AGTGTTCAGC TGTCATGGAC CCCAGGCGAT GACAACAATA 2040
GCCCCATTAC AACAATTCAT GACGAATATG AAGATGCAAT GCACAAGCCA GGGCTGTGGC 2100
ACCACCAAAC TGAAGTTTCT GGAACACAGA CCACAGCCCA GCTGAAGCTG TCTCCTTACG 2160
TGAACTACTC CTTCCGCGTG ATGGCAGTGA ACAGCATTGG GAAGAGCTTG CCCAGCGAGG 2220
CCTCTGAGCA GTATTTGACG AAAGCCTCAG AACCAGATAA AAACCCCACA GCTGTGGAAG 2280
GACTGGGATC AGAGCCTGAT AATTTGGTGA TTACGTGGAA GCCCTTGAAT GGTTTCGAAT 2340
TTAATGGGCC AGGCCTTCAG TACAAAGTTA GCTGGCGCCA GAAAGTTGGT GATGATGAAT 2400
GGACATCTGT GGTTGTGGCA AATGTATCCA AATATATTGT TTCAGGCACG CCAACCTTTG 2460
TTCCATACCT GATCAAAGTT CAGGCCCTGA ATGACATGGG GTTTGCCCCC GAGCCAGCTG 2520
TAGTCATGGG ACATTCTGGA GAAGACCTCC CAATGGTGGC TCCTGGGAAC GTGCGTGTGA 2580
ATGTGGTGAA CAGTACCTTA GCCGAGGTGC ACTGGGACCC AGTACCTCTG AAAAGCATCC 2640
GAGGACACCT ACAAGGCTAT CGGATTTACT ATTGGAAGAC CCAGAGTTCA TCTAAAAGAA 2700
ACAGACGTCA CATTGAGAAA AAGATCCTCA CCTTCCAAGG CAGCAAGACT CATGGCATGT 2760
TGCCGGGGCT AGAGCCCTTT AGCCACTACA CACTGAATGT CCGAGTGGTC AATGGGAAAG 2820
GGGAGGGCCC AGCCAGCCCT GACAGAGTCT TTAATACTCC AGAAGGAGTC CCCAGCGTTC 2880
CCTCGTCTTT GAAGATTGTG AATCCAACAC TGGACTCTCT CACTTTGGAA TGGGATCCAC 2940
CGAGCCACCC GAATGGCATT TTGACAGAGT ACACCTTAAA GTATCAGCCA ATTAACAACA 3000
CACATGAATT AGGCCCTCTG GTAGATTTGA AAATTCCTGC CAACAAGACA CGGTGGACTT 3060
TAAAAAATTT AAATTTCACC ACTCGATATA AGTTTTATTT CTATGCACAA ACATCAGCAG 3120
GATCAGGAAG TCAAATTACA GAGGAAGCAG TAACAACTGT GGATGAAGCT GGTATTCTTC 3180
CACCTGATGT AGGTGCAGGC AAAGTTCAAG CAGTAAATCC CAGGATCAGC AATCTTACTG 3240
CTGCAGCTGC TGAAACCTAT GCCAATATCA GTTGGGAATA TGAGGGACCA GAGTATGCCA 3300
ACTTTTATGT TGAATATGGT GTAGCAGGCA GCAAAGAAGA ATGGAGAAAA GAAATTGTAA 3360
ATGGTTCTCG GAGCTTCTTT GGGTTAAAGG GTCTAATGCC AGGAACAGCA TACAAGTTTC 3420

GAGTTGGTGC TGTGGGGGGA CCCCGGTTTG TGAGTTCAGA GGGTGTGTTT GAGACAGGCC 3480
CAGCGATGGC AAGCCGGCAG GTGGATATTG CAACTCAGGG CTGGTTCATT GGTCTGATGT 3540
GTGCTGTTGC TCTCCTTATC TTAATTTTGC TGATTGTTTG CTTCATCAGA AGAAACAAGG 3600
GTGGTAAATA TCCAGTTAAA GAAAAGGAAG ATGCCCATGC TGACCCTGAA ATCCAGCCTA 3660
TGAAGGAAGA TGATGGGACA TTTGGAGAAT ACAGTGATGC AGAAGACCAC AAGCCTTTGA 3720
AAAAAGGAAG TCGAACTCCT TCAGACAGGA CTGTGAAAAA AGAAGATAGT GACGACAGCC 3780
TACTTGACTA TGGAGAAGGG GTTAATGGCC AGTTCAATGA GGATGGCTCC TTTATTGGAC 3840
AATACAGTGG TAAAAAAGAG AAAGAGCCGG CTGAAGGAAA CGAAAGCTCA GAGGCACCTT 3900
CTCCTGTCAA CGCCATGAAT TCCTTTGTTT AATCATAGAA CTTGATTCCG ATGATGTCTT 3960
TACAGTTTGT TTGCTATTGT CCATCCAGGT TGTACTG 3997

INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1304 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gln Leu Lys Ile Met Pro Lys Lys Lys Arg Leu Ser Ala Gly Arg
1 5 10 15
Val Pro Leu Ile Leu Phe Leu Cys Gln Met Ile Ser Ala Leu Glu Val
20 25 30
Pro Leu Asp Pro Lys Leu Leu Glu Asp Leu Val Gln Pro Pro Thr Ile
35 40 45
Thr Gln Gln Ser Pro Lys Asp Tyr Ile Ile Asp Pro Arg Glu Asn Ile
50 55 60
Val Ile Gln Cys Glu Ala Lys Gly Lys Pro Pro Pro Ser Phe Ser Trp
65 70 75 80
Thr Arg Asn Gly Thr His Phe Asp Ile Asp Lys Asp Pro Leu Val Thr
85 90 95
Met Lys Pro Gly Thr Gly Thr Leu Ile Ile Asn Ile Met Ser Glu Gly

100 105 110

Lys Ala Glu Thr Tyr Glu Gly Val Tyr Gln Cys Thr Ala Arg Asn Glu

115 120 125

Arg Gly Ala Ala Val Ser Asn Asn Ile Val Val Arg Pro Ser Arg Ser

130 135 140

Pro Leu Trp Thr Lys Glu Lys Leu Glu Pro Ile Thr Leu Gln Ser Gly

145 150 155 160

Gln Ser Leu Val Leu Pro Cys Arg Pro Pro Ile Gly Leu Pro Pro Pro

165 170 175

Ile Ile Phe Trp Met Asp Asn Ser Phe Gln Arg Leu Pro Gln Ser Glu

180 185 190

Arg Val Ser Gln Gly Leu Asn Gly Asp Leu Tyr Phe Ser Asn Val Leu

195 200 205

Pro Glu Asp Thr Arg Glu Asp Tyr Ile Cys Tyr Ala Arg Phe Asn His

210 215 220

Thr Gln Thr Ile Gln Gln Lys Gln Pro Ile Ser Val Lys Val Ile Ser

225 230 235 240

Val Asp Glu Leu Asn Asp Thr Ile Ala Ala Asn Leu Ser Asp Thr Glu

245 250 255

Phe Tyr Gly Ala Lys Ser Ser Arg Glu Arg Pro Pro Thr Phe Leu Thr

260 265 270

Pro Glu Gly Asn Ala Ser Asn Lys Glu Glu Leu Arg Gly Asn Val Leu

275 280 285

Ser Leu Glu Cys Ile Ala Glu Gly Leu Pro Thr Pro Ile Ile Tyr Trp

290 295 300

Ala Lys Glu Asp Gly Met Leu Pro Lys Asn Arg Thr Val Tyr Lys Asn

305 310 315 320

Phe Glu Lys Thr Leu Gln Ile Ile His Val Ser Glu Ala Asp Ser Gly

325 330 335

Asn Tyr Gln Cys Ile Ala Lys Asn Ala Leu Gly Ala Ile His His Thr

340 345 350

Ile Ser Val Arg Val Lys Ala Ala Pro Tyr Trp Ile Thr Ala Pro Gln

355 360 365

Asn Leu Val Leu Ser Pro Gly Glu Asp Gly Thr Leu Ile Cys Arg Ala

370 375 380

Asn Gly Asn Pro Lys Pro Arg Ile Ser Trp Leu Thr Asn Gly Val Pro

385 390 395 400
Ile Glu Ile Ala Pro Asp Asp Pro Ser Arg Lys Ile Asp Gly Asp Thr
405 410 415
Ile Ile Phe Ser Asn Val Gln Glu Arg Ser Ser Ala Val Tyr Gln Cys
420 425 430
Asn Ala Ser Asn Glu Tyr Gly Tyr Leu Leu Ala Asn Ala Phe Val Asn
435 440 445
Val Leu Ala Glu Pro Pro Arg Ile Leu Thr Pro Ala Asn Thr Leu Tyr
450 455 460
Gln Val Ile Ala Asn Arg Pro Ala Leu Leu Asp Cys Ala Phe Phe Gly
465 470 475 480
Ser Pro Leu Pro Thr Ile Gln Trp Phe Lys Gly Ala Lys Gly Ser Ala
485 490 495
Leu His Glu Asp Ile Tyr Val Leu His Glu Asn Gly Thr Leu Glu Ile
500 505 510
Pro Val Ala Gln Lys Asp Ser Thr Gly Thr Tyr Thr Cys Val Ala Arg
515 520 525
Asn Lys Leu Gly Met Ala Lys Asn Glu Val His Leu Glu Ile Lys Asp
530 535 540
Pro Thr Trp Ile Val Lys Gln Pro Glu Tyr Ala Val Val Gln Arg Gly
545 550 555 560
Ser Met Val Ser Phe Glu Cys Lys Val Lys His Asp His Thr Leu Ser
565 570 575
Leu Thr Val Leu Trp Leu Lys Asp Asn Arg Glu Leu Pro Ser Asp Glu
580 585 590
Arg Phe Thr Val Asp Lys Asp His Leu Val Val Ala Asp Val Ser Asp
595 600 605
Asp Asp Ser Gly Thr Tyr Thr Cys Val Ala Asn Thr Thr Leu Asp Ser
610 615 620
Val Ser Ala Ser Ala Val Leu Ser Val Val Ala Pro Thr Pro Thr Pro
625 630 635 640
Ala Pro Val Tyr Asp Val Pro Asn Pro Pro Leu Asp Leu Glu Leu Thr
645 650 655
Asp Gln Leu Asp Lys Ser Val Gln Leu Ser Trp Thr Pro Gly Asp Asp
660 665 670
Asn Asn Ser Pro Ile Thr Thr Ile His Asp Glu Tyr Glu Asp Ala Met

675 680 685

His Lys Pro Gly Leu Trp His His Gln Thr Glu Val Ser Gly Thr Gln

690 695 700

Thr Thr Ala Gln Leu Lys Leu Ser Pro Tyr Val Asn Tyr Ser Phe Arg

705 710 715 720

Val Met Ala Val Asn Ser Ile Gly Lys Ser Leu Pro Ser Glu Ala Ser

725 730 735

Glu Gln Tyr Leu Thr Lys Ala Ser Glu Pro Asp Lys Asn Pro Thr Ala

740 745 750

Val Glu Gly Leu Gly Ser Glu Pro Asp Asn Leu Val Ile Thr Trp Lys

755 760 765

Pro Leu Asn Gly Phe Glu Phe Asn Gly Pro Gly Leu Gln Tyr Lys Val

770 775 780

Ser Trp Arg Gln Lys Val Gly Asp Asp Glu Trp Thr Ser Val Val Val

785 790 795 800

Ala Asn Val Ser Lys Tyr Ile Val Ser Gly Thr Pro Thr Phe Val Pro

805 810 815

Tyr Leu Ile Lys Val Gln Ala Leu Asn Asp Met Gly Phe Ala Pro Glu

820 825 830

Pro Ala Val Val Met Gly His Ser Gly Glu Asp Leu Pro Met Val Ala

835 840 845

Pro Gly Asn Val Arg Val Asn Val Val Asn Ser Thr Leu Ala Glu Val

850 855 860

His Trp Asp Pro Val Pro Leu Lys Ser Ile Arg Gly His Leu Gln Gly

865 870 875 880

Tyr Arg Ile Tyr Tyr Trp Lys Thr Gln Ser Ser Ser Lys Arg Asn Arg

885 890 895

Arg His Ile Glu Lys Lys Ile Leu Thr Phe Gln Gly Ser Lys Thr His

900 905 910

Gly Met Leu Pro Gly Leu Glu Pro Phe Ser His Tyr Thr Leu Asn Val

915 920 925

Arg Val Val Asn Gly Lys Gly Glu Gly Pro Ala Ser Pro Asp Arg Val

930 935 940

Phe Asn Thr Pro Glu Gly Val Pro Ser Val Pro Ser Ser Leu Lys Ile

945 950 955 960

Val Asn Pro Thr Leu Asp Ser Leu Thr Leu Glu Trp Asp Pro Pro Ser

965 970 975

His Pro Asn Gly Ile Leu Thr Glu Tyr Thr Leu Lys Tyr Gln Pro Ile

980 985 990

Asn Asn Thr His Glu Leu Gly Pro Leu Val Asp Leu Lys Ile Pro Ala

995 1000 1005

Asn Lys Thr Arg Trp Thr Leu Lys Asn Leu Asn Phe Thr Thr Arg Tyr

1010 1015 1020

Lys Phe Tyr Phe Tyr Ala Gln Thr Ser Ala Gly Ser Gly Ser Gln Ile

1025 1030 1035 104

Thr Glu Glu Ala Val Thr Thr Val Asp Glu Ala Gly Ile Leu Pro Pro

1045 1050 1055

Asp Val Gly Ala Gly Lys Val Gln Ala Val Asn Pro Arg Ile Ser Asn

1060 1065 1070

Leu Thr Ala Ala Ala Ala Glu Thr Tyr Ala Asn Ile Ser Trp Glu Tyr

1075 1080 1085

Glu Gly Pro Glu Tyr Ala Asn Phe Tyr Val Glu Tyr Gly Val Ala Gly

1090 1095 1100

Ser Lys Glu Glu Trp Arg Lys Glu Ile Val Asn Gly Ser Arg Ser Phe

1105 1110 1115 112

Phe Gly Leu Lys Gly Leu Met Pro Gly Thr Ala Tyr Lys Phe Arg Val

1125 1130 1135

Gly Ala Val Gly Gly Pro Arg Phe Val Ser Ser Glu Gly Val Phe Glu

1140 1145 1150

Thr Gly Pro Ala Met Ala Ser Arg Gln Val Asp Ile Ala Thr Gln Gly

1155 1160 1165

Trp Phe Ile Gly Leu Met Cys Ala Val Ala Leu Leu Ile Leu Ile Leu

1170 1175 1180

Leu Ile Val Cys Phe Ile Arg Arg Asn Lys Gly Gly Lys Tyr Pro Val

1185 1190 1195 120

Lys Glu Lys Glu Asp Ala His Ala Asp Pro Glu Ile Gln Pro Met Lys

1205 1210 1215

Glu Asp Asp Gly Thr Phe Gly Glu Tyr Ser Asp Ala Glu Asp His Lys

1220 1225 1230

Pro Leu Lys Lys Gly Ser Arg Thr Pro Ser Asp Arg Thr Val Lys Lys

1235 1240 1245

Glu Asp Ser Asp Asp Ser Leu Leu Asp Tyr Gly Glu Gly Val Asn Gly

1250 1255 1260

Gln Phe Asn Glu Asp Gly Ser Phe Ile Gly Gln Tyr Ser Gly Lys Lys

1265 1270 1275 128

Glu Lys Glu Pro Ala Glu Gly Asn Glu Ser Ser Glu Ala Pro Ser Pro

1285 1290 1295

Val Asn Ala Met Asn Ser Phe Val

1300

What is claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence encoding the NrCAMvar polypeptide of SEQ ID:NO2; or a nucleotide sequence complementary to said nucleotide sequence.

2. The polynucleotide of claim 1 which is DNA or RNA.

3. The polynucleotide of claim 2 wherein said nucleotide sequence comprises the NrCAMvar polypeptide encoding sequence contained in SEQ ID:NO2.

4. The polynucleotide of SEQ ID NO: 1.

5. A polynucleotide probe or primer comprising at least 15 contiguous nucleotides of the polynucleotide of claim 3.

6. A DNA or RNA molecule comprising an expression system, wherein said expression system is capable of producing a NrCAMvar polypeptide comprising an amino acid sequence of SEQ ID:NO2 when said expression system is present in a compatible host cell.

7. A host cell comprising the expression system of claim 7.

8. A process for producing a NrCAMvar polypeptide comprising culturing a host of claim 7 and under conditions sufficient for the production of said polypeptide.

9. The process of claim 8 which further includes recovering the polypeptide from the culture.

10. A process for producing a cell which produces a NrCAMvar polypeptide thereof comprising transforming or transfecting a host cell with the expression system of claim 6 such that the host cell, under appropriate culture conditions, produces a NrCAMvar polypeptide.

11. Cells produced by the process of claim 10.

12. A polypeptide which comprises the amino acid sequence of SEQ ID NO:2.

13. The polypeptide encoded in SEQ ID:NO2.

14. A NrCAMvar polypeptide prepared by the method of claim 9.

15. An antibody immunospecific for the NrCAMvar polypeptide of claim 12.

16. A method for the treatment of a subject in need of enhanced NrCAMvar polypeptide activity comprising:

(a) administering to the subject a therapeutically effective amount of an agonist to said polypeptide; and/or

(b) providing to the subject NrCAMvar polynucleotide in a form so as to effect production of said polypeptide activity *in vivo*.

17. A method for the treatment of a subject having need to inhibit NrCAMvar polypeptide activity comprising:

(a) administering to the subject a therapeutically effective amount of an antagonist to said polypeptide; and/or

(b) administering to the subject a nucleic acid molecule that inhibits the expression of the nucleotide sequence encoding said polypeptide; and/or

(c) administering to the subject a therapeutically effective amount of a polypeptide that competes with said polypeptide for its ligand, substrate , or receptor.

18. A process for diagnosing a disease or a susceptibility to a disease in a subject related to expression or activity of NrCAMvar polypeptide in a subject comprising:

(a) determining the presence or absence of a mutation in the nucleotide sequence encoding said NrCAMvar polypeptide in the genome of said subject; and/or

(b) analyzing for the presence or amount of the NrCAMvar polypeptide expression in a sample derived from said subject.

19. A method for identifying compounds which inhibit (antagonize) or agonize the NrCAMvar polypeptide which comprises:

(a) contacting a candidate compound with cells which express the NrCAMvar polypeptide (or cell membrane expressing NrCAMvar polypeptide) or respond to NrCAMvar polypeptide; and

(b) observing the binding, or stimulation or inhibition of a functional response; or comparing the ability of the cells (or cell membrane) which were contacted with the candidate compounds with the same cells which were not contacted for NrCAMvar polypeptide activity.

20. An agonist identified by the method of claim 19.

21. An antagonist identified by the method of claim 19.

22. A polynucleotide consisting essentially of a DNA sequence obtainable by screening an appropriate library containing the NrCAMvar gene under stringent hybridization conditions with a probe having the sequence of SEQ ID:NO 1 or a fragment thereof; and isolating said DNA sequence.

23. A polypeptide obtainable by expressing a nucleotide sequence comprising that of SEQ ID NO:1

24. A method for the treatment of diabetes, obesity or cancer which comprises administering to the subject a therapeutically effective amount of a modulator of NrCAMvar polypeptide activity.

25. A process according to claim 20 for diagnosing presence of or susceptibility to diabetes, obesity or cancer.

Figure 1 DNA sequence and amino acid sequence of NrCAMvar (SEQ ID NOS : 1 and 2 respectively

SEQ ID NO:1

CCATCGTAATTCGCCTAATGCAGCTTAAAATAATGCCGAAAAAGAAGCGCTTATCTGCGG
GCAGAGTGCCCCTGATTCTCTTCCTGTgCCAGATGATTAGTGCACTGGAAGTACCTCTTG
ATCCAAAACTTCTTGAAGACTTGGTACAGCCTCCAACCATCACCCAACAGTctccAAAAG
ATTACATTATTGACCCTCGGGAGAATATTGTAATCCAGTGTGAAGCCAAAGGGAAACCGC
CCCCAAGCTTTTCCTGGACCCGTAATGGGACTCATTTTGACATCGATAAAGACCCTCTGG
TCACCATGAAGCCTGGCACAGGAACGCTCATAATTAACATCATGAGCGAAGGGAAAGCTG
AGACCTATGAAGGAGTCTATCAGTGTACAGCAAGGAACGAACGCGGAGCTGCAGTTTCTA
ATAACATTGTTGTCCGCCCATCCAGATCACCATTGTGGACCAAAGAAAAACTTGAACCAA
TCACACTTCAAAGTGGTCAGTCTTTAGTACTTCCCTGCAGACCCCCAATTGGATTACCAC
CACCTATAATATTTTGGATGGATAATTCCTTTCAAAGACTTCCACAAAGTGAGAGAGTTT
CTCAAGGTTTGAATGGGGACCTTTATTTTTCCAATGTCCTCCCAGAGGACACCCGCGAAG
ACTATATCTGTTATGCTAGATTTAATCATACTCAAACCATACAGCAGAAGCAACCTATTT
CTGTGAAGGTGATTTCAGTGGATGAATTGAATGACACTATAGCTGCTAATTTGAGTGACA
CTGAGTTTTATGGTGCTAAATCAAGTAGAGAGAGGCCACCAACATTTTTAACTCCAGAAG
GCAATGCAAGTAACAAAGAGGAATTAAGAGGAAATGTGCTTTCACTGGAGTGCATTGCAG
AAGGACTGCCTACCCCAATTATTTACTGGGCAAAGGAAGATGGAATGCTACCCAAAAACA
GGACAGTTTATAAGAACTTTGAGAAAACCTTGCAGATCATTCATGTTTCAGAAGCAGACT
CTGGAAATTACCAATGTATAGCAAAAAATGCATTAGGAGCCATCCACCATACCATTTCTG
TTAGAGTTAAAGCGGCTCCATACTGGATCACAGCCCCTCAAAATCTTGTGCTGTCCCCAG
GAGAGGATGGGACCTTGATCTGCAGAGCTAATGGCAACCCCAAACCCAGAATTAGCTGGT
TAACAAATGGAGTCCCAATAGAAATTGCCCCTGATGACCCCAGCAGAAAAATAGATGGCG
ATACCATTATTTTTTCAAATGTTCAAGAAAGATCAAGTGCAGTATATCAGTGCAATGCCT
CTAATGAATATGGATATTTACTGGCAAACGCATTTGTAAATGTGCTGGCTGAGCCACCAC
GAATCCTCACACCTGCAAACACACTCTACCAGGTCATTGCAAACAGGCCTGCTTTACTAG
ACTGTGCCTTCTTTGGGTCACCTCTCCCAACCATCCAGTGGTTTAAAGGAGCTAAAGGAA
GTGCTCTTCATGAAGATATTTATGTTTTACATGAAAATGGAACTTTGGAAATTCCTGTGG
CCCAAAAGGACAGTACAGGAACTTATACGTGTGTTGCAAGGAATAAATTAGGGATGGCGA
AGAATGAAGTTCACTTAGAAATCAAAGATCCTACATGGATCGTTAAACAGCCCGAATATG
CAGTTGTGCAAAgAGGGAGCATGGTGTCCTTTGAATGCAAAGTGAAACATGATCACACCT
TATCCCTCACTGTCCTGTGGCTGAAGGACAACAGGGAACTGCCCAGTGATGAAAGGTTCA
CTGTTGACAAGGATCATCTAGTGGTAGCTGATGTCAGTGACGATGACAGCGGGACCTACA

CTCCTACTCCAACTCCAGCTCCCGTTTACGATGTCCCAAATCCTCCGCTTGACTTAGAAC
TGACAGATCAACTTGACAAAAGTGTTCAGCTGTCATGGACCCCAGGCGATGACAACAATA
GCCCCATTACAACAATTCATGACGAATATGAAGATGCAATGCACAAGCCAGGGCTGTGGC
ACCACCAAACTGAAGTTTCTGGAACACAGACCACAGCCCAGCTGAAGCTGTCTCCTTACG
TGAACTACTCCTTCCGCGTGATGGCAGTGAACAGCATTGGGAAGAGCTTGCCCAGCGAGG
CCTCTGAGCAGTATTTGACGAAAGCCTCAGAACCAGATAAAAACCCCACAGCTGTGGAAG
GACTGGGATCAGAGCCTGATAATTTGGTGATTACGTGGAAGCCCTTGAATGGTTTCGAAT
TTAATGGGCCAGGCCTTCAGTACAAAGTTAGCTGGCGCCAGAAAGTTGGTGATGATGAAT
GGACATCTGTGGTTGTGGCAAATGTATCCAAATATATTGTTTCAGGCACGCCAACCTTTG
TTCCATACCTGATCAAAGTTCAGGCCCTGAATGACATGGGGTTTGCCCCCGAGCCAGCTG
TAGTCATGGGACATTCTGGAGAAGACCTCCCAATGGTGGCTCCTGGGAACGTGCGTGTGA
ATGTGGTGAACAGTACCTTAGCCGAGGTGCACTGGGACCCAGTACCTCTGAAAAGCATCC
GAGGACACCTACAAGGCTATCGGATTTACTATTGGAAGACCCAGAGTTCATCTAAAAGAA
ACAGACGTCACATTGAGAAAAAGATCCTCACCTTCCAAGGCAGCAAGACTCATGGCATGT
TGCCGGGGCTAGAGCCCTTTAGCCACTACACACTGAATGTCCGAGTGGTCAATGGGAAAG
GGGAGGGCCCAGCCAGCCCTGACAGAGTCTTTAATACTCCAGAAGGAGTCCCCAGCGTTC
CCTCGTCTTTGAAGATTGTGAATCCAACACTGGACTCTCTCACTTTGGAATGGGATCCAC
CGAGCCACCCGAATGGCATTTTGACAGAGTACACCTTAAAGTATCAGCCAATTAACAACA
CACATGAATTAGGCCCTCTGGTAGATTTGAAAATTCCTGCCAACAAGACACGGTGGACTT
TAAAAAATTTAAATTTCACCACTCGATATAAGTTTTATTTCTATGCACAAACATCAGCAG
GATCAGGAAGTCAAATTACAGAGGAAGCAGTAACAACTGTGGATGAAGCTGGTATTCTTC
CACCTGATGTAGGTGCAGGCAAAGTTCAAGCAGTAAATCCCAGGATCAGCAATCTTACTG
CTGCAGCTGCTGAAACCTATGCCAATATCAGTTGGGAATATGAGGGACCAGAGTATGCCA
ACTTTTATGTTGAATATGGTGTAGCAGGCAGCAAAGAAGAATGGAGAAAAGAAATTGTAA
ATGGTTCTCGGAGCTTCTTTGGGTTAAAGGGTCTAATGCCAGGAACAGCATACAAGTTTC
GAGTTGGTGCTGTGGGGGGACCCCGGTTTGTGAGTTCAGAGGGTGTGTTTGAGACAGGCC
CAGCGATGGCAAGCCGGCAGGTGGATATTGCAACTCAGGGCTGGTTCATTGGTCTGATGT
GTGCTGTTGCTCTCCTTATCTTAATTTTGCTGATTGTTTGCTTCATCAGAAGAAACAAGG
GTGGTAAATATCCAGTTAAAGAAAAGGAAGATGCCCATGCTGACCCTGAAATCCAGCCTA
TGAAGGAAGATGATGGGACATTTGGAGAATACAGTGATGCAGAAGACCACAAGCCTTTGA
AAAAAGGAAGTCGAACTCCTTCAGACAGGACTGTGAAAAAAGAAGATAGTGACGACAGCC
TACTTGACTATGGAGAAGGGGTTAATGGCCAGTTCAATGAGGATGGCTCCTTTATTGGAC
AATACAGTGGTAAAAAAGAGAAAGAGCCGGCTGAAGGAAACGAAAGCTCAGAGGCACCTT
CTCCTGTCAACGCCATGAATTCCTTTGTTTAATCATAGAACTTGATTCCGATGATGTCTT
TACAGTTTGTTTGCTATTGTCCATCCAGGTTGTACTG

SEQ ID NO:2

MQLKIMPKKKRLSAGRVPLILFLCQMISALEVPLDPKLLEDLVQPPTITQQSPKDYIIDP
RENIVIQCEAKGKPPPSFSWTRNGTHFDIDKDPLVTMKPGTGTLIINIMSEGKAETYEGV
YQCTARNERGAAVSNNIVVRPSRSPLWTKEKLEPITLQSGQSLVLPCRPPIGLPPPIIFW
MDNSFQRLPQSERVSQGLNGDLYFSNVLPEDTREDYICYARFNHTQTIQQKQPISVKVIS
VDELNDTIAANLSDTEFYGAKSSRERPPTFLTPEGNASNKEELRGNVLSLECIAEGLPTP
IIYWAKEDGMLPKNRTVYKNFEKTLQIIHVSEADSGNYQCIAKNALGAIHHTISVRVKAA
PYWITAPQNLVLSPGEDGTLICRANGNPKPRISWLTNGVPIEIAPDDPSRKIDGDTIIFS
NVQERSSAVYQCNASNEYGYLLANAFVNVLAEPPRILTPANTLYQVIANRPALLDCAFFG
SPLPTIQWFKGAKGSALHEDIYVLHENGTLEIPVAQKDSTGTYTCVARNKLGMAKNEVHL
EIKDPTWIVKQPEYAVVQRGSMVSFECKVKHDHTLSLTVLWLKDNRELPSDERFTVDKDH
LVVADVSDDDSGTYTCVANTTLDSVSASAVLSVVAPTPTPAPVYDVPNPPLDLELTDQLD
KSVQLSWTPGDDNNSPITTIHDEYEDAMHKPGLWHHQTEVSGTQTTAQLKLSPYVNYSFR
VMAVNSIGKSLPSEASEQYLTKASEPDKNPTAVEGLGSEPDNLVITWKPLNGFEFNGPGL
QYKVSWRQKVGDDEWTSVVVANVSKYIVSGTPTFVPYLIKVQALNDMGFAPEPAVVMGHS
GEDLPMVAPGNVRVNVVNSTLAEVHWDPVPLKSIRGHLQGYRIYYWKTQSSSKRNRRHIE
KKILTFQGSKTHGMLPGLEPFSHYTLNVRVVNGKGEGPASPDRVFNTPEGVPSVPSSLKI
VNPTLDSLTLEWDPPSHPNGILTEYTLKYQPINNTHELGPLVDLKIPANKTRWTLKNLNF
TTRYKFYFYAQTSAGSGSQITEEAVTTVDEAGILPPDVGAGKVQAVNPRISNLTAAAAET
YANISWEYEGPEYANFYVEYGVAGSKEEWRKEIVNGSRSFFGLKGLMPGTAYKFRVGAVG
GPRFVSSEGVFETGPAMASRQVDIATQGWFIGLMCAVALLILILLIVCFIRRNKGGKYPV
KEKEDAHADPEIQPMKEDDGTFGEYSDAEDHKPLKKGSRTPSDRTVKKEDSDDSLLDYGE
GVNGQFNEDGSFIGQYSGKKEKEPAEGNESSEAPSPVNAMNSFV

Figure 2. Comparison of human NrCAM, NrCAMvar and chick NrCAM. The sequences compared are those of the current invention (Human1), of Lane *et al* (Human2) and chick. The nucleotide numbering refers to that of the Human1 sequence.

```
     1512                                                    AE10K1
Human1  GAAGATATTT ATGTTTTACA TGAAAATGGA ACTTTG---- ----------
Human2  GAAGATATTT ATGTTTTACA TGAAAATGGA ACTTTGGAAA TCAAAGATGC
Chick   GGAAATGAAT ATGTTTTCCA TGATAATGGA ACCTTG---- ----------

Human1  ---------- ------GAAA TTCCTGTGGC CCAAAAGGAC AGTACAGGAA
Human2  TACATGGATC GTTAAAGAAA TTCCTGTGGC CCAAAAGGAC AGTACAGGAA
Chick   ---------- ------GAAA TTCCAGTGGC TCAGAAGGAT AGTACTGGCA

Human1  CTTATACGTG TGTTGCAAGG AATAAATTAG GGATGGCGAA GAATGAAGTT
Human2  CTTATACGTG TGTTGCAAGG AATAAATTAG GGATGGCAAA GAATGAAGTT
Chick   CATACACATG TGTTGCAAGG AATAAATTAG GGAAGACGCA AAATGAAGTA
                          AE10K2
Human1  CACTTAGAAA TCAAAGATCC TACATGGATC GTTAAACAGC CCGAATATGC
Human2  CACTTA---- ---------- ---------- ------CAGC CCGAATATGC
Chick   CAACTGGAAG TTAAAGACCC AACGATGATA ATTAAACAGC CACAGTACAA
```

```
Human1  AGTTGTGCAA AGAGGGAGCA TGGTGTCCTT TGAATGCAAA GTGAAACATG
Human2  AGTTGTGCAA AGAGGGAGCA TGGTGTCCTT TGAATGCAAA GTGAAACATG
Chick   AGTGATTCAG AGATCTGCCC AGGCTTCATT TGAGTGTGTA ATAAAACATG

Human1  ATCACACCTT ATCCCTCACT GTCCTGTGGC TGAAGGACAA CAGGGAACTG
HUman2  ATCACACCTT ATCCCTCACT GTCCTGTGGC TGAAGGACAA CAGGGAACTG
Chick   ATCCTACCTT AATACCAACA GTTATATGGC TGAAAGACAA TAATGAACTA

Human1  CCCAGTGATG AAAGGTTCAC TGTTGACAAG GATCATCTAG TGGTAGCTGA
Human2  CCCAGTGATG AAAGGTTCAC TGTTGACAAG GATCATCTAG TGGTAGCTGA
Chick   CCAGATGATG AAAGGTTTCT AGTTGGTAAA GACAACTTGA CCATTATGAA

Human1  TGTCAGTGAC GATGACAGCG GGACCTACAC GTGTGTGGCC AACACCACTC
Humna2  TGTCAGTGAC GATGACAGCG GGACCTACAC GTGTGTGGCC AACACCACTC
Chick   TGTAACTGAT AAAGATGATG GAACATATAC TTGCATAGTT AATACTACTC

Human1  TGGACAGCGT CTCCGCCAGC GCTGTGCTTA GCGTTGTTGC TCCTACTCCA
Human2  TGGACAGCGT CTCCGCCAGC GCTGTGCTTA GCGTTGTTGC TCCTACTCCA
Chick   TGGACAGTGT TTCAGCAAGT GCTGTGCTTA CTGTTGTTGC TGCTCCCCCA
          AE10                  1956
Human1  ACTCCAGCTC CCGTTTACGA TGTC
Human2  ACTCCAGCTC CCGTTTACGA TGTC
Chick   ACTCCAGCTA TCATTTACGC TCGG
```

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 98/00434

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/705 C12Q1/68 C07K16/28 A61K31/70
A61K35/00 A61K38/00 A61K38/17 A61K48/00 G01N33/53
G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K C12Q A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LANE R.P. ET AL.: "Characterization of a highly conserved human homolog to the chicken neural cell surface protein Bravo/Nr-CAM that maps to chromosome band 7q31" GENOMICS, vol. 35, no. 3, 1 August 1996, pages 456-465, XP002066828 cited in the application see abstract	5,15,18, 22,25
Y	see page 458 - page 459; figure 1 see page 461, left-hand column, paragraph 2 - page 464, left-hand column; figure 3 --- -/--	1-4, 6-14,16, 17, 19-21,23

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 June 1998	03.07.98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Macchia, G

3

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Internal Application No
PCT/GB 98/00434

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DAVIS J.Q. ET AL.: "Molecular composition of the node of Ranvier: identification of Ankyrin-binding cell adhesion molecules Neurofascin (Mucin+/third FNIII domain-) and NrCAM at nodal axon segments" THE JOURNAL OF CELL BIOLOGY, vol. 135, no. 5, December 1996, pages 1355-1367, XP002066829	5,15,22
Y	see page 1357, right-hand column, paragraph 3 see page 1362; figure 7	1-4, 6-14,23

Y	WO 96 32959 A (ACORDA THERAPEUTICS) 24 October 1996 see page 4, line 20 - page 6, line 25 see page 9, line 6 - page 10, line 14 see page 36, line 24 - page 40, line 31 see page 76 - page 84; tables 1-3 see page 103 - page 107; claims	16,17, 19-21

X	VOLKMER H. ET AL.: "Neurofascin induces neurites by heterophilic interactions with axonal NrCAM while NrCAM requires F11 in the axonal surface to extend neurites" THE JOURNAL OF CELL BIOLOGY, vol. 135, no. 4, November 1996, pages 1059-1069, XP002066882 see page 1062, left-hand column, paragraph 2 - page 1063; figure 2	15,19-21

X	SUTER D.M. ET AL.: "Binding between the neural cell adhesion molecules Axonin-1 and Nr-CAM/Bravo is involved in neuron-glia interaction" THE JOURNAL OF CELL BIOLOGY, vol. 131, no. 4, November 1995, pages 1067-1081, XP002066830 see page 1074; figure 5 see page 1072, right-hand column see page 1075; figure 6	15,19-21

P,X	Database EMBL, entry HSC7NRCAM, Accession number AJ001057, 30 November 1997, 100% identity with Seq.ID:1 nt.33-3932 XP002066831 see the whole document	1-5, 12-14, 22,23

3

INTERNATIONAL SEARCH REPORT

In. .ational application No.
PCT/GB 98/00434

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 16, 17, 24 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter. ...nal Application No
PCT/GB 98/00434

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9632959 A	24-10-1996	AU 5557296 A	07-11-1996
		EP 0821591 A	04-02-1998
		NO 974811 A	19-12-1997
		PL 322947 A	02-03-1998